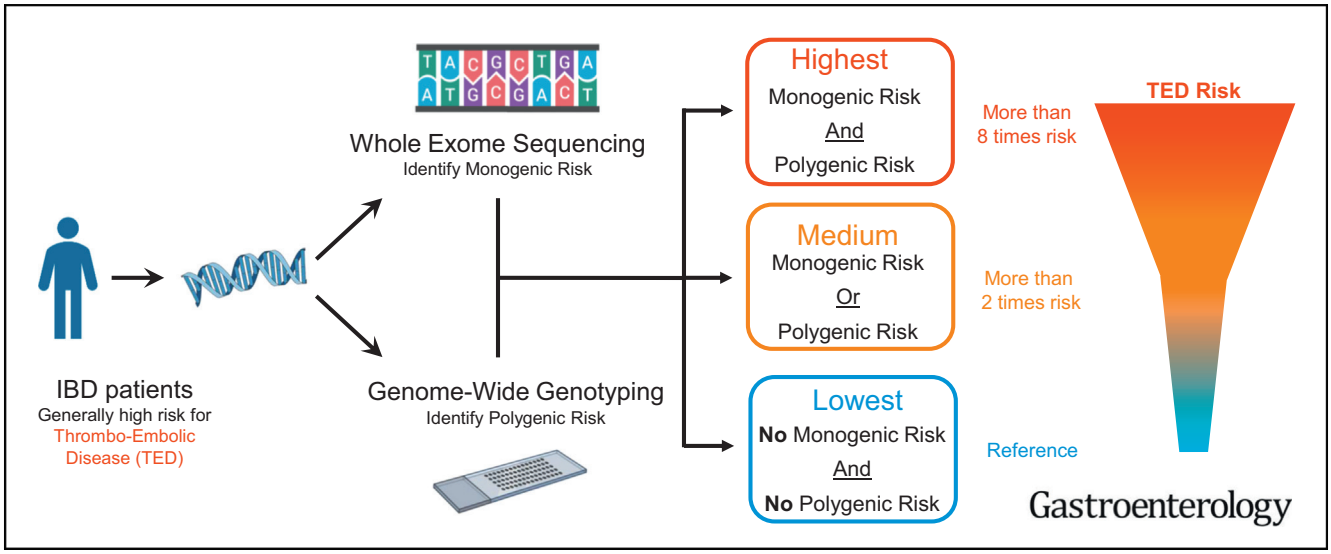
Prevalence and Effect of Genetic Risk of Thromboembolic Disease in Inflammatory Bowel Disease

Takeo Naito,[1,2] Gregory J. Botwin,[1] Talin Haritunians,[1] Dalin Li,[1] Shaohong Yang,[1] Michelle Khrom,[1] Jonathan Braun,[1] NIDDK IBD Genetics Consortium, Lisa Abbou,[1] Emebet Mengesha,[1] Christine Stevens,[3] Atsushi Masamune,[2] Mark Daly,[3] and Dermot P. B. McGovern[1]

[1]F. Widjaja Family Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, Los Angeles, California; [2]Division of Gastroenterology, Tohoku University Graduate School of Medicine, Sendai, Japan; and [3]Stanley Center for Psychiatric Research, Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts



See Covering the Cover synopsis on page 633.

BACKGROUND AND AIMS: The largest cause of mortality in patients with inflammatory bowel disease (IBD) remains thromboembolic disease (TED). Recent reports have demonstrated that both monogenic and polygenic factors contribute to TED and 10% of healthy subjects are genetically at high risk for TED. Our aim was to utilize whole-exome sequencing and genome-wide genotyping to determine the proportion of IBD patients genetically at risk for TED and investigate the effect of genetic risk of TED in IBD. **METHODS:** The TED polygenic risk score was calculated from genome-wide genotyping. Thrombophilia pathogenic variants were extracted from whole-exome sequencing. In total, 792 IBD patients had both whole-exome sequencing and genotyping data. We defined patients at genetically high risk for TED if they had a high TED polygenic risk score or carried at least 1 thrombophilia pathogenic variant. **RESULTS:** We identified 122 of 792 IBD patients (15.4%) as genetically high risk for TED. Among 715 of 792 subjects whose documented TED status were available, 63 of the 715 patients (8.8%) had TED events. Genetic TED risk was significantly associated with increased TED event (odds ratio, 2.5; $P = .0036$). In addition, we confirmed an additive effect of monogenic and polygenic risk on TED ($P = .0048$). Patients with high TED genetic risk more frequently had thrombosis at multiple sites (78% vs 42%, odds ratio, 3.96; $P = .048$). **CONCLUSIONS:** Genetic risk (both poly- and monogenic) was significantly associated with TED history. Our results suggest that genetic traits identify approximately 1 in 7 patients with IBD who will experience 2.5-fold or greater risk for TED.

Keywords: Genetics; Inflammatory Bowel Diseases; Thrombosis.

The inflammatory bowel diseases (IBDs) Crohn's disease (CD) and ulcerative colitis (UC) are chronic relapsing inflammatory conditions of the gastrointestinal tract.[1] The incidence of IBD is increasing, with a prevalence

Abbreviations used in this paper: **CD, Crohn's disease; IBD, inflammatory bowel disease; OR, odds ratio; PRS, polygenic risk score; QC, quality control; SNP, single nucleotide polymorphism; TED, thromboembolic disease; TPV, thrombophilia pathogenic variant; UC, ulcerative colitis; WES, whole-exome sequencing.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2020.10.019

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT:

Patients with inflammatory bowel disease (IBD) are at high risk of Thrombo-embolic disease (TED); however, prevalence and effect of genetic risk of TED in IBD remains unknown.

NEW FINDINGS:

By aggregating whole exome sequencing (WES) and whole-genome genotyping data, we identified that ~1 in 7 IBD patients are genetically at around 2.5 times higher risk for TED. Genetic risk was significantly related with increased risk of TED events and there was an additive effect of monogenic and polygenic risks.

LIMITATIONS:

Number of cases in our study was limited (N = 792). Thrombo-embolic events were determined retrospectively by chart review blinded to genetic status.

IMPACT:

We found that more than 15 % of IBD patients are at higher risk of TED. Genetic test by combining monogenic and polygenic risk can identify IBD patients with higher risk of TED. WES provides a more comprehensive evaluation of genetic risk in IBD.

of >1.3% in the United States[2] and a global prevalence that surpasses 0.3%.[3] Patients with IBD are reported to have 3- to 4-fold increased risk of thromboembolic disease (TED), and the most significant cause of mortality in IBD remains TED.[4] This increased risk appears to be unique to IBD, as other chronic inflammatory diseases, such as rheumatoid arthritis and celiac disease, do not confer this risk.[5] More recently, TED also has been identified as a potential complication of JAK inhibition, a mechanism of action recently approved for treating UC patients, which is now under a US Food and Drug Administration boxed warning for increased risk of thrombosis-associated morbidity and mortality. Due to the impact of TED on IBD prognosis and therapeutic management, developing methods to identify IBD patients at high risk for TED is an urgent clinical issue.

A number of factors can influence increased TED risk, including disease activity, hospitalization, age, pregnancy, medications, surgery, and genetics.[6–10] Previous reports for genetic risks of TED in IBD patients have focused mainly on monogenic variants, such as Factor V Leiden deficiency.[10,11] However, a recent study reported that risk of TED in the general population is influenced not only by monogenic risk, which classifies individuals based on the presence or absence of discrete large effect variants, but also polygenic risk, which refers to the effect of numerous loci, often of small individual effect, across the genome. In a study containing 2 very large population cohorts, a combination of polygenic risk score (PRS, which aggregates multiple genetic TED risk variants) with 2 monogenic variants (Factor V Leiden deficiency and prothrombin G20210A mutation) delineated 10% of individuals with an approximately 2.5-fold increased likelihood of developing TED compared with non–high-risk controls.[12] However, the prevalence and effects of polygenic and monogenic risk of TED in IBD patients have not been studied previously.

As the genetic etiology of TED becomes increasingly understood and the cost of genotyping/sequencing continues to decrease, the assessment of TED risk through genetics is becoming a viable clinical tool. If IBD patients with elevated risk of TED can be identified through genetic testing, there exists the potential to optimize drug therapy and the precise risks and benefits of anticoagulation prophylaxis can be considered on a personalized basis. To address this potential, the present study used whole-exome sequencing (WES) to assess monogenic risk, together with genome-wide genotyping of data to determine the proportion of IBD patients genetically at elevated risk for TED and investigate the effect of this risk on thrombotic events in IBD patients.

Methods

Study Design and Subjects

The study design is summarized in Figure 1. Samples were genotyped as part of the National Institute of Diabetes and Digestive and Kidney Disease Inflammatory Bowel Disease Genetics Consortium genotyping efforts. There were 11,584 samples available after stringent quality control (QC) (details below), of which 2452 samples were recruited at Cedars-Sinai Medical Center. The WES cohort consists of 3198 subjects recruited at Cedars-Sinai Medical Center and 340 subjects provided to Cedars-Sinai Medical Center by the National Laboratory for the Genetics of Israeli Populations; all samples were sequenced at the Broad Institute (see below). Admixture[13] analysis was used to calculate ethnicity proportion estimations for all individuals. Only subjects identified by admixture as European ancestry (EUR proportion $\geq$ 0.70) were included in the further analyses.

Seven hundred and ninety-two Cedars-Sinai European ancestry IBD cases had both whole-genome genotyping and WES data available. IBD was defined on the basis of clinical symptoms as well as standard endoscopic, radiographic, and histologic findings. Detailed clinical data, including patient sex, age at diagnosis, disease location and behavior (according to the Montreal Classification), surgical history, family history of IBD, and TED history were available in 715 cases and was collected from the medical records by 2 phenotypers (S.Y. and L.A.) who were blinded to the patients' genotype information. The Institutional Review Board of Cedars-Sinai Medical Center approved the study, and all patients provided written informed consent.

Global Screening Array Genotyping, Quality Control, and Imputation

Samples were genotyped on the Illumina Global Screening Array at Feinstein Institute for Medical Research or at the Broad Institute in Massachusetts. Pre-imputation single nucleotide polymorphism (SNP) QC metrics were applied for 700,078 SNPs, including exclusion of nonautosomal markers and variants with minor allele frequency <1%, genotyping missingness >3%, and deviation from Hardy-Weinberg equilibrium in controls $P < 1 \times 10^{-6}$. A total of 11,584 samples

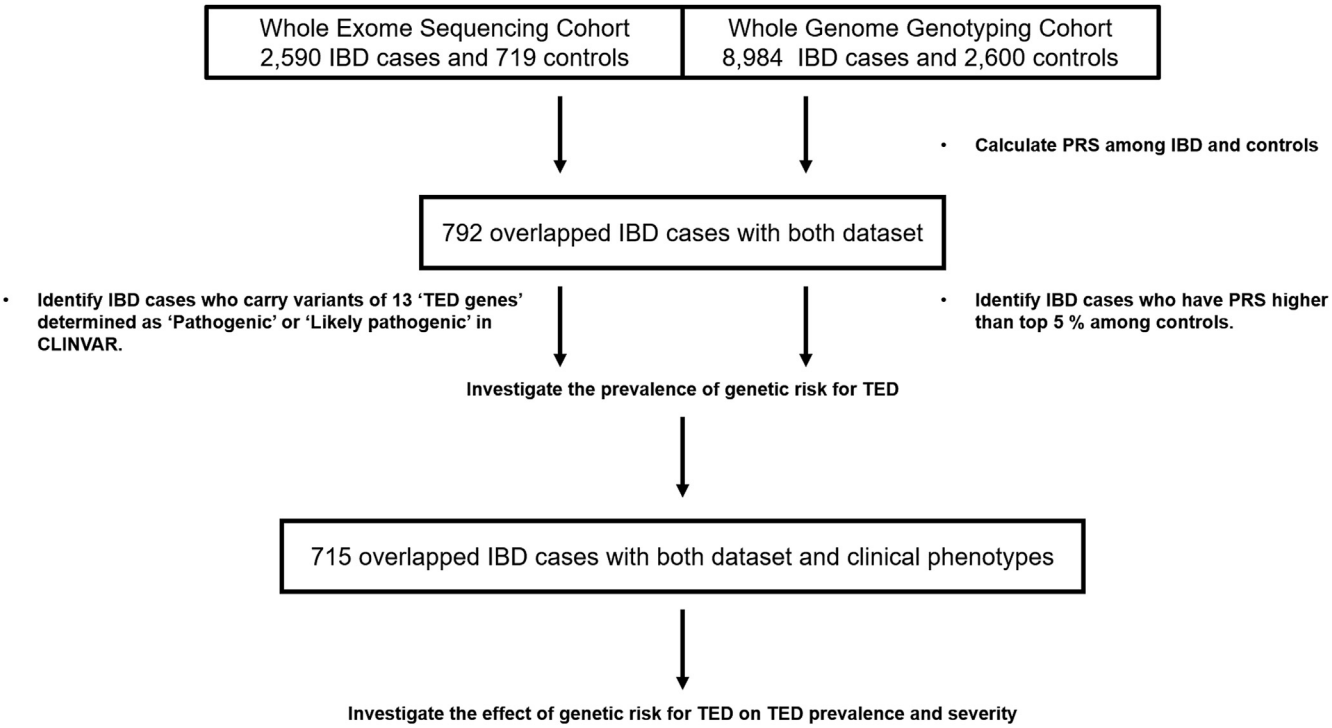


Figure 1. Study design and cohorts. Flow of study and outline of study cohorts. Detailed information is described in the Methods section.

(8984 IBD and 2600 control) passed sample QC, which included exclusion of samples with genotyping call rate <95%, sex discrepancies, duplicated samples, EUR proportion <0.70, ambiguous disease information, and sample permission use restrictions. Genotypes were phased using Eagle, version 2.3,[14] and imputation was performed using the Michigan Imputation Server[15] per instructions and HRC r1.1[16] reference panel. Variants with estimated imputation accuracy (Rsq) <0.3 were excluded post-imputation before analyses. In total, 10,357,915 SNPs passed variant QC.

Polygenic Risk Score Calculation and Normalization

Among 297 variants reported in the TED PRS,[12] 265 variants were available in our imputed Global Screening Array cohort. Using these 265 variants we generated the TED PRS using PLINK software, version 2.00a,[17] for 2600 controls and 8984 IBD cases. We normalized PRS into *Z* score and defined the top 5% normalized PRS among healthy control group as threshold.

Whole-Exome Sequencing

Paired-end WES was performed based on Illumina platform with 20X reading depth in 3538 subjects. Reads alignment to the human reference genome GRCh37 were performed using BWA and variant calling were performed based on GATK best practices. Individual variants with genotyping quality <65, depth <20, Strand odds ratio >3, or call rate <95% were removed. For SNPs, variants with ReadPosRankSum < −4 or Fisher Strand filter >60 were also removed. For indels, variants with ReadPosRankSum < −20 or Fisher Strand filter >200 were also removed. In total, 3,349,656 variants passed QC. Samples with a mean genotype quality <65, a depth <25, a genotype rate <96.5%, or a transition to transversion ratio <2.5 were removed from further analysis. Individuals of ambiguous disease information were removed. Individuals of ambiguous imputed sex or imputed sex inconsistent with reported sex were also removed. A total of 3309 samples (2590 IBD and 719 controls) passed QC.

Thrombophilia Pathogenic Variants Extraction

Using CLINVAR "pathogenic" or "likely pathogenic" classification,[18] we extracted variants located within 15 blood clotting–related genes, yielding a total of 7 different thrombophilia pathogenic variants (TPVs) (Supplementary Table 1 and Table 2). All QC and variant annotation was performed using Hail (Hail Team; Hail 0.2.36-ed011219dd93; https://github.com/hail-is/hail/releases/tag/0.2.36).

Definitions

We defined patients at high genetic risk for TED if they had a TED PRS more than the top 5% of the control population distribution or carried at least 1 TPV. This aligns with the report that individuals within the top 5% of the TED PRS have an approximately 2.5-fold increased risk of VTE relative to the rest of the population, which is similar to the risk attributed to the presence of monogenic variants.[12] Disease activity at the time of TED for CD was measured by the Harvey-Bradshaw Index and colonoscopy report at the time of clotting event (when available). Patients were considered to have active disease if they had Harvey-Bradshaw Index

Table 1. Univariate and Multivariate Models of Associations of Thromboembolic Disease History

Variable	TED		Univariate *P* value	Multivariate *P* value	OR	95% CI
	No	Yes				
N	652	63	—	—	—	—
Disease type, n (%)			—	—	—	—
CD	511 (78.37)	46 (73.02)	.3408			
UC	129 (19.79)	16 (25.4)	.3240			
IBDU	12 (1.84)	1 (1.59)	1.0000			
Sex, n (%)				—	—	—
Male	349 (53.53)	32 (50.79)	.6935			
Female	303 (46.47)	31 (49.21)	—			
Age at IBD onset, *y*, mean ± SD	23.77 ± 12.74	30.11 ± 17.83	.0075	.0001	1.04	1.02–1.06
Disease duration, *y*, mean ± SD	19.83 ± 11.65	23.74 ± 13.87	.0340	.0083	1.03	1.01–1.05
A1, n (%)	173 (34.06)	12 (26.09)	.1515			
A2, n (%)	287 (56.5)	27 (58.7)	.6424			
A3, n (%)	48 (9.45)	7 (15.22)	.3119			
Subjects missing age at onset data, n	3	0	—	—	—	—
Disease location, n (%)				—	—	—
L1	66 (13.15)	2 (4.35)	.1015			
L2	61 (12.15)	5 (10.87)	1.0000			
L3	375 (74.7)	39 (84.78)	.1121			
L4	85 (16.63)	7 (15.22)	1.0000			
Subjects missing disease location data, n	9	0		—	—	—
Disease behavior, n (%)				—	—	—
B1	183 (36.31)	12 (26.09)	.2008			
B2	253 (50.2)	27 (58.7)	.2340			
B3	192 (38.1)	25 (54.35)	.0600			
Subjects missing disease behavior data, n	7	0		—	—	—
Perianal disease, n (%)	296 (57.93)	28 (60.87)	.7565	—	—	—
Disease extent, n (%)				—	—	—
E1	12 (9.68)	0 (0)	.3620			
E2	41 (33.06)	4 (25)	.7762			
E3	71 (57.26)	12 (75)	.1811			
Subjects missing disease extent data, n	5	0		—	—	—
Extensive disease (E3 or L3) , n (%)	446 (69.91)	51 (80.95)	.1296	—	—	—
Smoking status, n (%)				—	—	—
Current smoking	141 (22.63)	16 (26.67)	.5203			
Past smoking	38 (6.1)	5 (8.33)	.4148			
Never smoking	444 (71.27)	39 (65)	.3024			
Subjects missing smoking data, n	29	3		—	—	—
IBD family history, n (%)	200 (31.6)	13 (21.67)	.1425	—	—	—
Subjects missing IBD family history data, n	19	3		—	—	—
Surgery history, n (%)				—	—	—
IBD-related bowel surgery	375 (57.52)	48 (76.19)	.0045	.0120	2.24	1.19–4.20
CD-related bowel surgery	333 (65.17)	38 (82.61)	.0147	—	—	—
Colectomy in UC	38 (29.46)	10 (62.5)	.0116	—	—	—
Genetic risk (high)	91 (13.96)	18 (28.57)	.0050	.0036	2.5	1.35–4.64

CI, confidence interval; IBDU, Inflammatory bowel disease undetermined.

Table 2. Characteristics of Inflammatory Bowel Disease Patients With Thromboembolic Disease Events

Characteristic	Low risk[a]	High risk[b]	*P* value
Patients, n	45	18	—
Female, n (%)	21 (46.67)	10 (55.56)	.5850
Male, n (%)	24 (53.33)	8 (44.44)	
Interval from IBD onset to TED, *y*, mean ± SD	18.91 ± 13.97	11.62 ± 14.21	.0863
Age at TED, *y*, mean ± SD	45.45 ± 17.24	50.58 ± 18.84	.3441
Active disease at TED, n (%)	25 (71.43)	10 (83.33)	.7027
Subjects missing disease activity data, n	10	6	
Hospitalization within 3 mo before TED, n (%)	14 (41.18)	8 (61.54)	.3279
Subjects missing hospitalizations data, n	11	5	
Current smoking, n (%)	13 (30.95)	3 (16.67)	.5224
Past smoking, n (%)	4 (9.52)	1 (5.56)	1.0000
Never smoking, n (%)	25 (59.52)	14 (77.78)	.1516
Subjects missing smoking data, n	3	0	
IBD-related surgery within 6 mo and before TED, n (%)	2 (5.41)	0 (0)	1.0000
Subjects missing surgery data, n	8	3	
Women taking OCP at time of TED,[c] n (%)	4 (22.22)	0 (0)	.2800
Subjects missing OCP data, n	3	1	
Biologic[d] use at the time of TED, n (%)	16 (37.21)	12 (70.59)	.0244
Subjects missing biologic use data, n	2	1	
Multiple sites of TED, n (%)	19 (45.24)	14 (77.78)	.0447
Subjects missing multiple sites data, n	3	0	
No. of sites, mean ± SD	2.15 ± 2.23	3.33 ± 3.44	.2343

OCP, oral contraceptive pills.
[a]Genetically low risk for TED.
[b]Genetically high risk for TED.
[c]Percentage of OCP represents percentage within women.
[d]Infliximab, adalimumab, or certolizumab.

scores ≥5 and/or endoscopy showed active disease, Disease activity at the time of TED for UC was evaluated by the full Mayo score. A full Mayo score >2 was considered as active disease. Smoking status was defined as currently smoking, past smoker, or never smoker, as assessed at the most recent clinical visit. Patients who had history of IBD surgery were defined as having history of IBD-related surgery (colectomy for UC and any bowel resection for CD) from their disease onset to last time of follow-up. Hospitalization for any reason within 3 months before TED event and use of oral contraceptive pills at the time of TED event were investigated. The use of any biologic therapy at the time of TED event was also investigated. Patients with TED were defined as patients who had a history of venous thrombosis at any site identified by ultrasonography or computed tomography scanning. If patients had multiple episodes of TED events in their disease course, we considered the first time of TED in terms of time of TED event.

Statistical Analysis

Fisher exact test (2-sided) was used to explore associations of categorical data in Tables 1 and 2. Unpaired *t* test was used to explore quantitative data between 2 groups in Tables 1 and 2. For the multivariate model in Table 1 and Supplementary Table 4, logistic regression with all univariate risk factors with a *P* value <.05 were included together in the multivariable logistic regression model along with the first 2 principal components. Linear trend test in logistic regression with age at last visit, history of IBD surgery, and the first 2 principal components was performed for Figure 2. Logistic regression with age at last visit, history of IBD surgery, and the first 2 principal components was performed in Figure 3. Logistic regression was performed with the first 2 principal components in Supplementary Table 5 and Supplementary Table 6. A *P* value <.05 was considered statistically significant. All statistical analyses were performed with R software, version 3.6.1 (http://www.r-project.org/).

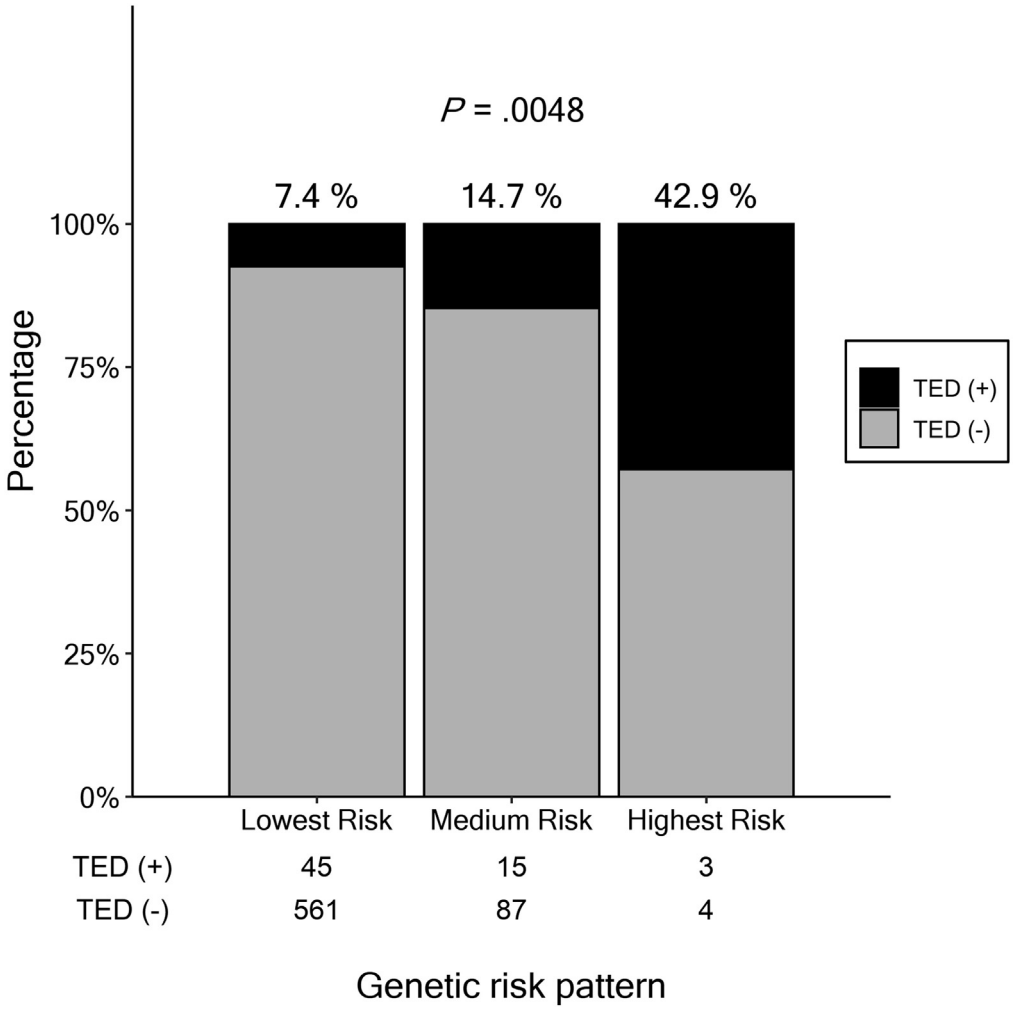


Figure 2. Percentage of TED events in each genetic risk group. The percentage of patients with TED in each group. X-axis shows classification of genetic risk for TED (lowest risk, patients without genetic risk; medium risk, patients who have 1 risk, either high PRS or carriage of a TPV; highest risk, patients who both high PRS and a TPV). The *numbers* below each bar represents number of TED positive and negative cases in each group. *P* value was calculated by trend test.

Results

Prevalence of Patients With Inflammatory Bowel Disease Who Are Genetically High Risk for Thromboembolic Disease

Among the 792 IBD subjects with both PRS and WES data, 49 had a high PRS and 82 carried at least 1 variant among the 7 identified TPVs, including Factor V Leiden and the prothrombin G20210A mutation (Supplementary Table 2 for full list of TPVs). In total, 122 of 792 IBD patients (15.4%) were identified as genetically high risk for TED.

Difference of Thromboembolic Disease Genetic Risk Among Patients With Inflammatory Bowel Disease and Controls

There was no difference in TED PRS distribution; frequency of Factor V Leiden mutation; and frequency of prothrombin G20210A mutation between 8984 IBD cases and 2600 controls (P = .84, .26, and .94, respectively; Supplementary Table 3).

Effect of Genetic Risk on Thromboembolic Disease Event in Multivariate Model

Among the 792 subjects with WES and genotyping data, detailed longitudinal clinical information on TED events was available for 715 subjects (109 high risk and 606 nonrisk patients). In total, 63 of the 715 patients (8.8%) had a documented TED event. Patients with TED had significantly longer disease duration (23.7 vs 19.8 years; P = .034), were older at IBD onset (30.1 vs 23.7 years; P = .0075), and were more likely to have had IBD surgery (odds ratio [OR], 2.36; P = .0045). No other demographic or clinical factors were statistically associated with TED (Table 1). After adjustment for age, disease duration, and history of IBD surgery, genetic TED risk was significantly associated with increased TED event (OR, 2.5; P = .0036) (Table 1). In addition, after adjusting multicollinearity between disease duration and age at disease onset, both high PRS and carriage of TPV were independently associated with TED, respectively (OR, 3.13; P = .0070 and OR. 2.11; P = .042) (Supplementary Table 4).

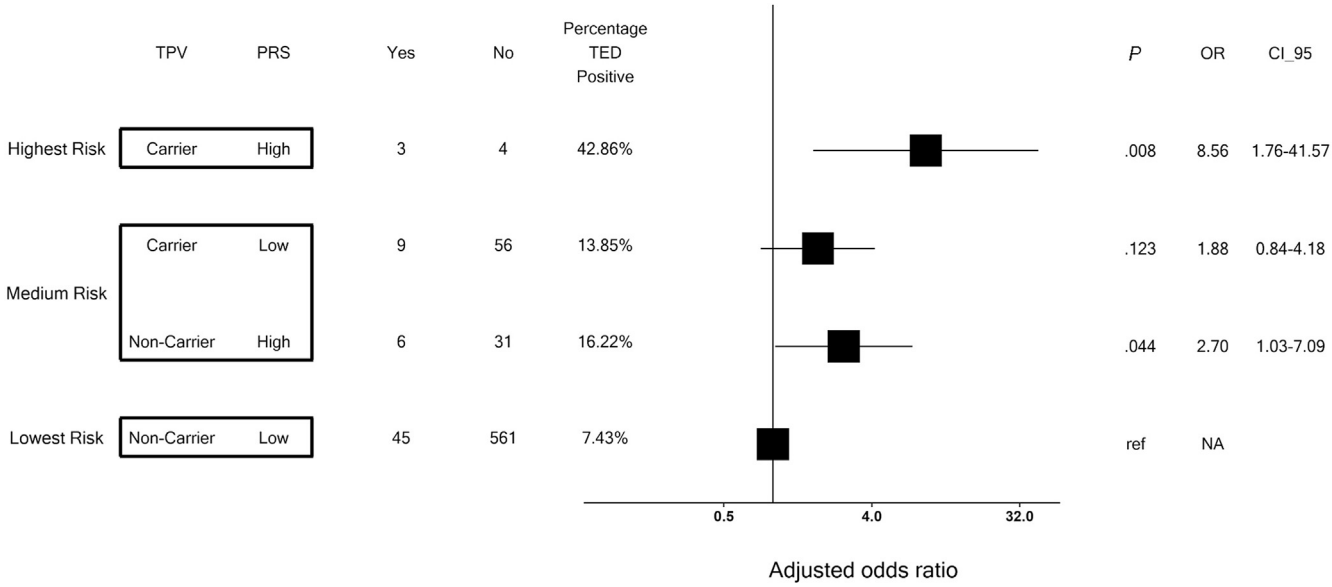


Figure 3. Risk of TED by monogenic and polygenic risk status. Patients were stratified into 2 groups according to their PRS—high or low defined as more than the top 5% of the control population distribution or the other 95%, respectively. For carriers and noncarriers of TPV in each PRS group, the *P* value and OR for TED were calculated in a logistic regression model with age at last visit and the first 2 principal components as covariates. Noncarriers with low PRS served as the reference group. Each *black square* represents mean of OR in each group and *horizontal line* around each square represents 95% confidence.

Additive Effect of Genetic Risk on Thromboembolic Disease Event in Inflammatory Bowel Disease

We subsequently confirmed an additive effect of genetic risk on TED; patients with both a high PRS and carrying pathogenic variants are at the highest risk of TED; patients who have 1 risk factor (either high PRS or carriage of a TPV) were medium risk; and patients without either of these genetic risk factors were at lowest risk (43%, 15%, and 7%, respectively, $P = .0048$; trend test) (Figure 2). Among patients with medium risk, patients with PRS had slightly higher risk than patients with TPV only (OR, 2.8 vs 1.8, respectively) (Figure 3).

Characteristic of Patients With Genetic Risk Within Thromboembolic Disease Cases

Patients with high TED genetic risk tended to have shorter time from IBD onset to TED event (11.62 vs 18.91 years; $P = .086$) and were more likely to have thrombosis documented at multiple sites (78% vs 42%; OR, 3.96; $P = .048$) (Table 2). Among these 2 groups, disease activity at the time of TED, hospitalization within 3 months before TED, and surgery before TED were not significantly different between the 2 groups. Patients with high TED genetic risk were more likely to have received biologics (71% vs 37%; OR, 3.95; $P = .024$). Biologics have previously been reported to have protective effect[19,20] on TED.

Discussion

By aggregating whole-genome genotyping and WES data, our analyses demonstrate that approximately 1 in 7 patients with IBD have odds 2.5 times higher than nongenetically high-risk patients with IBD for experiencing TED. Higher genetic risk was associated with TED events, suggesting that these IBD patients might warrant more aggressive prophylaxis against TED, and in whom JAK inhibitors might need to be used judiciously. TED PRS and TPV were independently associated with TED and also have additive effects on TED events. Furthermore, our analysis within TED cases suggests genetic risk also affects disease severity of TED. Although the number of patients who carry both PRS and TPV risks is small (7 patients), 3 developed TED at multiple sites and at a young age (mean, 27.5 years).

The prevalence of TED in our IBD cohort was 8.8%, consistent with previous findings.[21] The risk of TED in IBD patients is reported to be 3- to 4-fold higher compared with the general population,[6] which is attributed to risk factors, including disease flare, extended disease location, and steroid use.[22] Our study confirmed that long disease duration, older age at IBD onset, and history of IBD-related surgery are risk factors for TED. Disease duration and age at onset are strongly associated with age at last visit to hospital (Supplementary Figure 1*A* and *B*). If we include all 3 of these parameters into multivariate model, all of them became nonsignificant because of multicollinearity. Variance inflation factors of age at last visit, disease duration, and age at diagnosis in the model are 26.50, 13.65, and 17.11, respectively, which indicates strong multicollinearity among these variables (Supplementary Table 5). If we combined disease duration or age at IBD onset with age at last visit, only age at last visit remains significant (Supplementary Table 5). Age at last visit approximates current age, therefore, our study suggests older age, a well-established risk factor for TED, is the significant demographic risk factor for TED history.[23,24] We also confirmed that history of IBD-related surgery was

significantly associated with TED history (OR, 3.95 and 2.54 for colectomy history in UC and bowel resection for CD, respectively). Importantly, this effect was independent from time-dependent parameters. IBD-related surgery has previously been reported to be one of the established risk factors of TED in multiple studies, with a higher risk observed in UC cases requiring colectomy than CD-related surgery.[25–28] The elevated risk of UC-related surgery probably reflects, in part, the higher inflammatory burden, and perhaps this is a population in which an increased understanding of genetic risk might have the largest impact. Importantly, genetic factors remained significant even after adjusting for age at last visit and history of IBD-related surgery (Supplementary Table 4).

In the within-TED patients analyses (Table 2), we found that a high proportion of patients (74%) had active disease at the time of TED event, which is consistent with previous reports, including in a single-center retrospective study in which 71% of IBD patients were found to have active disease at the time of the TED event.[29] In addition, nearly half of patients (47%) had been hospitalized within 3 months before the TED event. These factors were not different between patients with or without genetic TED risk. Our results support prior evidence that active disease and hospitalization are risk factors of TED,[4,6,30] and demonstrate that these are independent from genetic risk. In our cohort, very few patients had undergone IBD-related surgery within 6 months before a TED event, were taking oral contraceptive pills at the time of TED (Table 2), and no one was receiving prophylactic warfarin or JAK inhibitor at the time of TED. Interestingly, patients with genetic risk of TED were receiving biologics more frequently than those without genetic risk at the time of TED, whereas disease activity was not different between the 2 groups. Generally, biologics have been reported to have a protective effect on TED,[19,20] suggesting that our findings might have underestimated the effect of genetic variation on TED risk. Finally, we found patients with genetic risk had multiple-site TEDs more frequently than those without genetic risk, which might indicate a more aggressive course of TED.

We did not find an elevated PRS TED distribution in IBD compared with controls. In addition, the frequency of the 2 major TPVs (Factor V Leiden and prothrombin G20210A mutation) were not elevated in IBD, consistent with previous studies.[10,21] Elevated genetic burden of TED risk does not alone explain the increased risk of TED in IBD patients (Supplementary Table 3).

The relationship between TPVs and disease behavior or extent in IBD has not been studied previously and any relationship is important for interpreting our findings. We observed no association between elevated genetic burden (carriage of TPV and high PRS) and disease behavior and extent of disease (Supplementary Table 6).

In our analyses, access to WES enabled us to also include rare TPVs beyond Factor V Leiden and prothrombin G20210A mutation. As shown in Supplementary Table 2, we identified 5 of 13 TED patients (38%) with "other" TPVs, suggesting that including only Factor V Leiden and prothrombin variants would miss more than one-third of "monogenic" TPVs. Looking at both common and rare TPVs is necessary for a more comprehensive estimate of monogenic TED risk. Furthermore, the additive effect we have shown here suggests that an accurate estimation of genetic risk requires both TPV and PRS assessment and that PRS might define a higher TED risk than TPVs alone.

With decreasing costs of sequencing and moves towards genomic medicine, it is likely that increasing numbers of people will have their genome sequenced, making "routine" integration of genetic risk assessment for various traits, including TED, possible. Our data suggest that approximately 1 in 7 IBD patients is at higher risk of TED and those IBD patients are at approximately 2.5 times increased risk of TED. Considering the relatively high prevalence of genetic risk and the fact that TED can lead to significant morbidity and even mortality, routine screening of TED genetic risk for patients with IBD can beneficially impact these poor outcomes. Additional studies are warranted, including perhaps randomized controlled trials of prophylactic anticoagulation in IBD patients at high risk of TED stratified by genetic risk. With likely increasing availability of these types of data, the IBD community requires the development of guidelines and strategies to determine whether these patients should be counseled about long-term TED prophylaxis and also whether drugs such as JAK inhibitors should be used cautiously or avoided in this setting.

Our study has limitations. First, because our study is a retrospective study there is some missing clinical information, such as corticosteroid use, indwelling catheters, and family history of TED. Therefore, it is difficult to estimate accurate effects of some clinical factors on TED. Second, because not all variants of TED PRS were available in our study, the result might be slightly different if all variants were included, although if this has had an effect it is likely to have underestimated the true genetic contribution to TED. Third, because our study is a single-center analysis, the number of subjects is relatively small and additional cohorts will need to be studied. In addition, as is the case with the majority of genetic studies currently, our study is limited to European ancestry subjects and future studies will be needed in other populations so that the benefits of precision medicine approaches, such as the one described, are available to all parts of society.

In conclusion, we have demonstrated that approximately 1 in 7 IBD patients are genetically at a higher risk for TED and genetic risk is independently associated with TED events when adjusting for time-dependent parameters and IBD-related surgery. For comprehensive prediction of genetic risk of TED, both monogenic and polygenic approaches are needed. To our knowledge, this is the first report suggesting benefits for clinical decision-making in IBD through combining both WES and whole-genome genotyping. With increased interest in genomic sequencing/genotyping for clinical utility, our findings suggest that strategies for managing patients at high risk of TED identified through genetic approaches should be developed.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2020.10.019.

References

1. Abraham C, Cho JH. Inflammatory bowel disease. N Engl J Med 2009;361:2066–2078.
2. Dahlhamer JM, Zammitti EP, Ward BW, et al. prevalence of inflammatory bowel disease among adults aged ≥18 years—United States, 2015. MMWR Morb Mortal Wkly Rep 2016;65:1166–1169.
3. **Ng SC**, Shi HY, **Kaplan G**, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2018;390:2769–2778.
4. Nguyen GC, Sam J. Rising prevalence of venous thromboembolism and its impact on mortality among hospitalized inflammatory bowel disease patients. Am J Gastroenterol 2008;103:2272–2280.
5. Miehsler W, Reinisch W, Valic E, et al. Is inflammatory bowel disease an independent and disease specific risk factor for thromboembolism? Gut 2004;53:542–548.
6. Grainge MJ, West J, Card TR. Venous thromboembolism during active disease and remission in inflammatory bowel disease: a cohort study. Lancet 2010;375:657–663.
7. Heit JA, Kobbervig CE, James AH, et al. Trends in the incidence of venous thromboembolism during pregnancy or postpartum: a 30-year population-based study. Ann Intern Med 2005;143:697–706.
8. Kappelman MD, Horvath-Puho E, Sandler RS, et al. Thromboembolic risk among Danish children and adults with inflammatory bowel diseases: a population-based nationwide study. Gut 2011;60:937–943.
9. Sarlos P, Szemes K, Hegyi P, et al. Steroid but not biological therapy elevates the risk of venous thromboembolic events in inflammatory bowel disease: a meta-analysis. J Crohns Colitis 2018;12:489–498.
10. Papa A, Danese S, Grillo A, et al. Review article: inherited thrombophilia in inflammatory bowel disease. Am J Gastroenterol 2003;98:1247–1251.
11. **Liang J, Wu S**, Feng B, et al. Factor V Leiden and inflammatory bowel disease: a systematic review and meta-analysis. J Gastroenterol 2011;46:1158–1166.
12. Klarin D, Busenkell E, Judy R, et al. Genome-wide association analysis of venous thromboembolism identifies new risk loci and genetic overlap with arterial vascular disease. Nature Genetics 2019;51:1574–1579.
13. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 2009;19:1655–1664.
14. Loh PR, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 2016;48:1443–1448.
15. **Das S, Forer L, Schönherr S**, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287.
16. **McCarthy S, Das S, Kretzschmar W**, et al. A reference panel of 64,976 haplotypes for genotype imputation 2016;48:1279–1283.
17. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
18. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res 2018;46:D1062–D1067.
19. deFonseka AM, Tuskey A, Conaway MR, et al. Antitumor necrosis factor-α therapy is associated with reduced risk of thromboembolic events in hospitalized patients with inflammatory bowel disease. J Clin Gastroenterol 2016;50:578–583.
20. Higgins PD, Skup M, Mulani PM, et al. Increased risk of venous thromboembolic events with corticosteroid vs biologic therapy for inflammatory bowel disease. Clin Gastroenterol Hepatol 2015;13:316–321.
21. Giannotta M, Tapete G, Emmi G, et al. Thrombosis in inflammatory bowel diseases: what's the link? Thromb J 2015;13:14.
22. Cheng K, Faye AS. Venous thromboembolism in inflammatory bowel disease. World J Gastroenterol 2020;26:1231–1241.
23. Faye AS, Wen T, Ananthakrishnan AN, et al. Acute venous thromboembolism risk highest within 60 days after discharge from the hospital in patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol 2020;18:1133–1141.e3.
24. McCurdy JD, Israel A, Hasan M, et al. A clinical predictive model for post-hospitalisation venous thromboembolism in patients with inflammatory bowel disease. Aliment Pharmacol Ther 2019;49:1493–1501.
25. Alhassan N, Trepanier M, Sabapathy C, et al. Risk factors for post-discharge venous thromboembolism in patients undergoing colorectal resection: a NSQIP analysis. Tech Coloproctol 2018;22:955–964.
26. McKechnie T, Wang J, Springer JE, et al. Extended thromboprophylaxis following colorectal surgery in patients with inflammatory bowel disease: a comprehensive systematic clinical review. Colorectal Dis 2020;22:663–678.
27. Wallaert JB, De Martino RR, Marsicovetere PS, et al. Venous thromboembolism after surgery for inflammatory bowel disease: are there modifiable risk factors? Data from. ACS NSQIP. Dis Colon Rectum 2012;55:1138–1144.
28. Alatri A, Schoepfer A, Fournier N, et al. Prevalence and risk factors for venous thromboembolic complications in the Swiss Inflammatory Bowel Disease Cohort. Scand J Gastroenterol 2016;51:1200–1205.
29. Bollen L, Vande Casteele N, Ballet V, et al. Thromboembolism as an important complication of inflammatory bowel disease. Eur J Gastroenterol Hepatol 2016;28:1–7.
30. Jackson E, Curtis KM, Gaffield ME. Risk of venous thromboembolism during the postpartum period: a systematic review. Obstet Gynecol 2011;117:691–703.

Author names in bold designate shared co-first authorship.

Received July 29, 2020. Accepted October 7, 2020.

Correspondence
Address correspondence to: Dermot P. B. McGovern, MD, PhD, F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, Cedars-Sinai Medical Center, 8730 Alden Drive, Los Angeles, California 90048. e-mail: dermot.mcgovern@cshs.org.

Acknowledgments
DNA samples for a subset of Jewish controls from Cedars-Sinai Medical Center were obtained from The National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University (http://yoran.tau.ac.il/nlgip/). This study was supported by the Cedars-Sinai MIRIAD IBD Biobank. The MIRIAD IBD Biobank is supported by the Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, National Institute of Diabetes and Digestive and Kidney Disease grants P01DK046763 and U01DK062413, and The Leona M. and Harry B. Helmsley Charitable Trust.

Date repository: Our original data including raw genetic data and metadata is available at github (https://github.com/mcgovernlab).

Members of the NIDDK IBD Genetics Consortium: Judy Cho; Mark Silverberg; Richard Duerr; John Rioux; Steven Brant; and Philip Schumm.

CRediT Authorship Contributions
Takeo Naito, MD, PhD (Data curation: Supporting; Formal analysis: Lead; Investigation: Lead; Methodology: Equal; Validation: Equal; Visualization: Lead; Writing – original draft: Lead)

Gregory J. Botwin, BASc (Data curation: Equal; Methodology: Equal; Supervision: Supporting; Validation: Supporting; Writing – review & editing: Equal)

Talin Haritunians, PhD (Data curation: Lead; Investigation: Supporting; Methodology: Equal; Validation: Supporting; Writing – review & editing: Supporting),

Dalin Li, MD, PhD (Data curation: Supporting; Formal analysis: Supporting; Methodology: Equal; Supervision: Equal; Validation: Lead; Writing – review & editing: Supporting),

Shaohong Yang, MD, PhD (Data curation: Equal; Methodology: Equal; Resources: Lead; Writing – review & editing: Supporting)

Michelle Khrom, BASc (Data curation: Equal; Formal analysis: Supporting; Methodology: Supporting; Writing – review & editing: Supporting)

Jonathan Braun, MD, PhD (Supervision: Equal; Writing – review & editing: Equal). NIDDK IBD Genetics Consortium, none (Data curation: Supporting)

Lisa Abbou, BASc (Data curation: Equal; Resources: Equal)

Emebet Mengesha, BASc (Data curation: Equal; Resources: Equal)

Christine Stevens, PhD (Data curation: Equal; Resources: Equal)

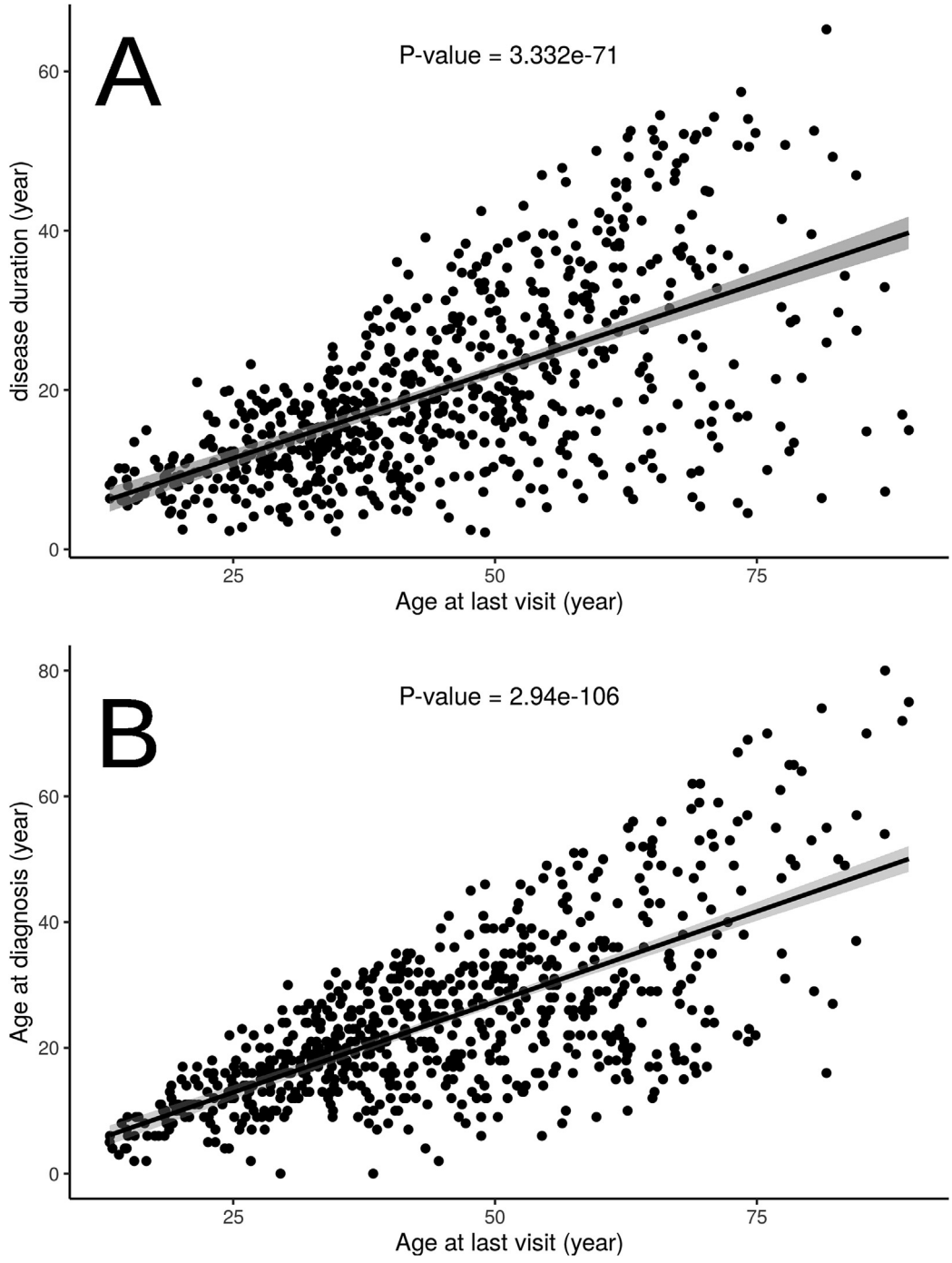
Atsushi Masamune, MD, PhD (Conceptualization: Supporting; Writing – review & editing: Supporting)

Mark Daly, PhD (Data curation: Equal; Resources: Equal)

Dermot P.B. McGovern, MD, PhD (Conceptualization: Lead; Formal analysis: Equal; Funding acquisition: Lead; Project administration: Lead; Supervision: Lead; Writing – original draft: Equal; Writing – review & editing: Lead)

Conflicts of interest
These authors disclose the following: Dermot P. B. McGovern , Talin Haritunians, Dalin Li, and Jonathan Braun are faculty members at Cedars-Sinai Medical Center. Takeo Naito, Michelle Khrom, Gregory J. Botwin, Shaohong Yang, Lisa Abbou, and Emebet Mengesha are employees at Cedars-Sinai. Cedars-Sinai has financial interests in Prometheus Biosciences, Inc, a company that has access to the data and specimens in Cedars-Sinai's MIRIAD Biobank (including the data and specimens used in this study). Prometheus Biosciences, Inc seeks to develop commercial products. Dermot P. B. McGovern and Dalin Li are paid consultants and shareholders of Prometheus Biosciences, Inc. Dermot P. B. McGovern is a consultant for Gilead Sciences, Boehringer-Ingelheim, Pfizer, Bridge Biotherapeutics, Qu Biologics, Prometheus Biosciences, Takeda, and Palatin Technologies and grant support from Janssen.

Funding
This work was supported by internal funds from the F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute. The Cedars-Sinai MIRIAD IBD Biobank is supported by the F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute, National Institutes of Health/National Institute of Diabetes and Digestive and Kidney Diseases (grants P01 DK046763 and U01 DK062413), and The Leona M and Harry B Helmsley Charitable Trust.



Supplementary Figure 1. Association of age at last visit with disease duration and age at diagnosis. (*A*) Association between age at last visit and disease duration. *X-axis* represents age at last visit and *Y-axis* represents disease duration. *P* value for association was calculated by linear regression without any covariates. *Blue line* represents regression line and *shaded area* represents 95% confidence. (*B*) Association between age at last visit and age at diagnosis. The form of figure is the same as described in (*A*).

Supplementary Table 1. List of Candidate Thromboembolic Disease–Related Genes

Gene	Associated disease
F5	Factor V Leiden
F2	Prothrombin mutation
PROC	Protein C deficiency
PROS1	Protein S deficiency
SERPINC1	Antithrombin III deficiency
THBD	Defect in thrombomodulin
ADAMTS13	Thrombotic thrombocytopenic purpura
MTR	Elevated homocysteine
APOA	Elevated lipoprotein(a)
FGA, FGB, FGG	Congenital dysfibrinogenemia
F8	Elevated factor VIII
HRG	Thrombophilia due to both elevated and deficient histidine-rich glycoprotein
KNG1	High-molecular-weight kininogen deficiency

Supplementary Table 2. Thrombophilia Pathogenic Variants Identified in Our Whole-Exome Sequencing Cohort

CHR	POS[a]	REF	ALT	Genes	Definition in CLINVAR[b]	Related disease in CLINVAR	No. of IBD subjects (total)	No. of IBD subjects with TED
1	169519049	T	C	*F5*	Drug response	Hormonal contraceptives for systemic use response, toxicity	37	5
1	169524537	C	G	*F5*	Pathogenic	Thrombophilia due to activated protein C resistance	2	1
1	173883881	G	A	*SERPINC1*	Pathogenic	Reduced antithrombin III activity/antithrombin deficiency	1	1
9	136291338	G	C	*ADAMTS13*	Likely pathogenic	Upshaw-Schulman syndrome	1	0
9	136302010	C	T	*ADAMTS13*	Likely pathogenic	Upshaw-Schulman syndrome	1	1
11	46747447	G	A	*F2*	Pathogenic	Prothrombin type 3	3	2
11	46761055	G	A	*F2*	Pathogenic	Venous thrombosis	40	3

ALT, alternative allele; CHR, chromosome; POS, position; REF, reference allele.
[a]Chromosomal positions are based on the Genome Reference Consortium human build 37 (GRCh37).
[b]Definition in CLINVAR is based on ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/vcf_GRCh37/archive_2.0/2019/clinvar_20190722.vcf.

Supplementary Table 3. Distribution of Thromboembolic Disease Polygenic Risk Score and Frequency of 2 major Thrombophilia Pathogenic Variants Among Patients With Inflammatory Bowel Disease and Controls

Variable	Control	IBD	*P* value[a]	β	SE
TED PRS, mean ± SD	−0.054 ± 0.999	0.015 ± 1	.835	.0048	0.0229
Factor V Leiden, %	2.747	2.287	.258	−.1144	0.1012
Prothrombin (*F2*) G20210A mutation, %	1.464	1.917	.937	−.0105	0.1329

[a]*P* values were calculated by logistic regression with 2 principal components.

Supplementary Table 4. Multivariate Analysis of Thromboembolic Disease History for Age at Last Visit, History of Inflammatory Bowel Disease Bowel Surgery, and Genetic Risk With Variance Inflation Factor[a]

Variable	*P* value	OR	95% CI	VIF
Results with genetic risk				
Genetic risk (high)[b]	.0035	2.51	1.35–4.65	1.01
History of IBD bowel surgery	.0201	2.08	1.12–3.87	1.03
Age at last visit	.00003	1.04	1.02–1.05	1.03
Results with separated genetic risk (TPV and PRS)				
PRS (high)[c]	.0070	3.13	1.37–7.18	1.01
TPV carrier[d]	.0433	2.11	1.02–4.34	1.01
History of IBD bowel surgery	.0241	2.04	1.10–3.79	1.03
Age at last visit	.00002	1.04	1.02–1.06	1.03

CI, confidence interval; VIF, variance inflation factor.
[a]*P* values, ORs, and VIFs were calculated by logistic regression with 2 principal components.
[b]Genetic risk (high) is defined as having a high TED polygenic risk score (more than the top 5% of the control population distribution) or carried at least 1 TPV.
[c]PRS (high) is defined as more than the top 5% of the control population distribution of TED PRS.
[d]TPV carrier is defined as having at least 1 TPV.

Supplementary Table 5. Multivariate Analysis of Thromboembolic Disease History for Time-Dependent Parameters, History of Inflammatory Bowel Disease Bowel Surgery and Genetic Risk With Variance Inflation Factor[a]

Variable	*P* value	OR	95% CI	VIF
Results with all time-dependent parameters (disease duration, age at diagnosis, and age at last visit)				
Genetic risk (high)[b]	.0035	2.52	1.35–4.68	1.01
History of IBD bowel surgery	.0159	2.18	1.16–4.09	1.06
Disease duration	.6043	0.98	0.90–1.06	13.65
Age at diagnosis	.7444	0.99	0.91–1.07	17.11
Age at last visit	.2227	1.06	0.97–1.15	26.50
Results with 2 time-dependent parameters (disease duration and age at last visit)				
Genetic risk (high)[b]	.0034	2.53	1.36–4.70	1.01
History of IBD bowel surgery	.0152	2.18	1.16–4.11	1.06
Disease duration	.4774	0.99	0.97–1.02	1.63
Age at last visit	.0001	1.04	1.02–1.06	1.57
Results with 2 time-dependent parameters (age at diagnosis and age at last visit)				
Genetic risk (high)[b]	.0034	2.52	1.36–4.69	1.01
History of IBD bowel surgery	.0164	2.17	1.15–4.07	1.06
Age at diagnosis	.5560	1.01	0.98–1.03	2.04
Age at last visit	.0065	1.03	1.01–1.06	2.10

CI, confidence interval; VIF, variance inflation factor.
[a]*P* values, ORs, and VIFs were calculated by logistic regression with 2 principal components.
[b]Genetic risk (high) is defined as having a high TED polygenic risk score (more than the top 5% of the control population distribution) or carried at least 1 TPV.

Supplementary Table 6. Association Between the Presence of Thrombophilia Pathogenic Variants and Polygenic Risk Score and History of Inflammatory Bowel Disease–Related Surgery and Extensive Disease[a]

Variable	Presence of TPVs, n (%)		*P* value	OR (95% CI)
	Carrier[b]	Noncarrier		
History of IBD-related surgery	43 (59.72)	380 (59.10)	.91	1.03 (0.62–1.69)
Extensive disease (E3 or L3)	54 (75.00)	443 (68.90)	.18	1.52 (0.82–2.82)
	PRS, n (%)			
	High[c]	Low		
History of IBD-related surgery	26 (59.09)	397 (59.17)	.92	0.97 (0.52–1.81)
Extensive disease (E3 or L3)	29 (65.91)	468 (69.75)	.27	0.70 (0.36–1.33)

CI, confidence interval.
[a]*P* values and ORs were calculated by logistic regression with 2 principal components.
[b]TPV carrier is defined as having at least 1 TPV.
[c]PRS high is defined as more than the top 5% of the control population distribution of TED PRS.